

INTEGRATING GIS IN A STATEWIDE MEDICAL EDUCATION  
ADMINISTRATIVE SYSTEM

Ashley Michelle Davis

Submitted to the faculty of the University Graduate School  
in partial fulfillment of the requirements  
for the degree  
Master of Science  
in the Department of Geography,  
Indiana University

April 2019

Accepted by the Graduate Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

Master's Thesis Committee

---

Jeffrey S. Wilson, PhD, MA, BS, Chair

---

Komal Kochhar, MBBS, MHA

---

Vijay O. Lulla, PhD, MA, BE

© 2019

Ashley Michelle Davis

## ACKNOWLEDGEMENTS

First, I would like to thank my thesis committee chair, Jeffrey S. Wilson, PhD. Thank you for providing me with this wonderful opportunity to expand my academic career. Thank you for your patience and encouragement; it has been a long process, but the knowledge and skills I have gained since starting this project are immeasurable. Secondly, thank you for my committee members: Komal Kochhar, MBBS and Vijay Lulla, Ph.D. Your support and feedback have been invaluable to this process.

I would like to thank my grandfather, James Moore, for always believing in me and instilling the value of education in me. He has taught me to be resilient and to always stay positive despite obstacles that may arise. He is a shining light in my life and I would not be the person I am today without him.

Lastly, I would like to thank my husband, Jordan Davis, for helping me through this project. If it was not for him, I would have never started this journey towards my MS GIS. His encouragement and belief in me kept me going whenever I was discouraged. Thank you for sitting with me in coffee shops for who-knows-how-many-hours to make sure I stayed on task. You're the best and now it's your turn.

Ashley Michelle Davis

INTEGRATING GIS IN A STATEWIDE MEDICAL EDUCATION  
ADMINISTRATIVE SYSTEM

Geographic technologies can be used to visualize and analyze data patterns that may go unnoticed from other approaches. The purpose of this project was to provide examples of how GIS and cartographic methods are being used to help facilitate communication and inform management processes for a complex statewide medical education system administered by the Indiana University School of Medicine, the largest medical school in the United States. The IU School of Medicine has nine regional campuses located around the state in addition to numerous partnering hospitals where medical students are trained. We illustrate geographic examples of various stages of medical student education from admissions, through campus assignments and clinical rotations, to residency training. These geographic processes are being used to inform reaccreditation processes as well as assisting administration with recruitment/retention strategies, statewide planning, and analysis in a complex medical education system.

Jeffrey S. Wilson, PhD, MA, BS, Chair

## TABLE OF CONTENTS

List of Figures .....	vii
List of Abbreviations .....	viii
Introduction.....	1
Background .....	6
Hometown/Practice Location .....	6
Specialty Area.....	8
Data .....	11
Methods.....	15
Results.....	20
Map Narratives .....	20
Admissions Maps.....	20
Indiana County Maps.....	20
United States Maps .....	24
Medical Student Education Maps .....	27
Graduate Medical Education World Map.....	31
Discussion .....	33
Conclusions.....	34
References .....	36
Curriculum Vitae .....	

## LIST OF FIGURES

Figure 1: Statewide distribution of IU School of Medicine and IU Health facilities in 2016.....	3
Figure 2: Hospitals and clinical location partners located in Indiana in 2016.....	4
Figure 3: Medical student education pipeline from applicant to resident/fellow. ....	11
Figure 4: Residents and fellows plans after completing their residency or fellowship training at IUSM based on GME databases.....	14
Figure 5: The Python code used to concatenate first and last name fields into one field .....	17
Figures 6a – 6f: Admissions maps identifying the number of IUSM applicants from Indiana counties during the 2010 - 2015 academic years. ....	21
Figure 7a -7f: Admissions maps displaying the number of IUSM applicants by state during the 2010 – 2015 academic years.....	24
Figure 8a – 8e: Medical Student Education (MSE) maps illustrating the total number of students by class year at each campus during 2010 – 2014 academic years .....	28
Figure 9: Percent of medical students matching into either primary care or non-primary care residency programs by first year campus location .....	30
Figure 10: The Graduate Medical Education (GME) world map, depicting the country of origin for graduates entering the residency or fellowship program at IUSM ..	32

## LIST OF ABBREVIATIONS

GIS: Geographic Information Systems  
MSE: Medical Student Education  
GME: Graduate Medical Education  
IUSM: Indiana University School of Medicine



## INTRODUCTION

The Indiana University School of Medicine (IUSM), founded in 1903, was recognized in 2015-16 as the largest medical school in the nation with over 1,400 MD students, 160 PhD students, and almost 500 students enrolled in other graduate degree programs (IUSM, 2016). Administration of the medical school is headquartered in Indianapolis on the Indiana University - Purdue University Indianapolis (IUPUI) campus. However, the school's medical education system is distributed across nine campuses throughout the state (Figure 1) and includes partnerships with about 50 different hospitals and clinical locations (Figure 2).

Given the size, complexity, and spatially distributed nature of the school's medical education system, the IUSM Office of Educational Affairs began a collaboration in 2015 with the IUPUI Department of Geography to explore the potential for integrating geographic technology in the management and analysis of its medical student data. This project was prompted by the recognition that using geographic technology to visualize and analyze spatial and temporal patterns derived from the school's diverse student databases has the potential to facilitate communication and insight that may not be readily apparent from other approaches. Geographic visualization and analytical approaches could be particularly useful in contexts such as communicating information to state legislatures, external evaluators, and administrators within the school, as well as evaluating the school's impact on the health workforce needs and service to medically underserved areas and health professional shortage areas in both the state and the nation.

To help facilitate this collaboration, the IUSM established a new internship opportunity for graduate students in the Masters of Science in Geographic Information

Science (MS GIS) program at IUPUI. Students worked closely under the supervision of faculty and professional staff from the IUPUI Department of Geography and the IUSM Office of Research in Medical Education. This thesis describes the initial results that grew out of this collaboration and directions this interdisciplinary team plans to pursue in future work.

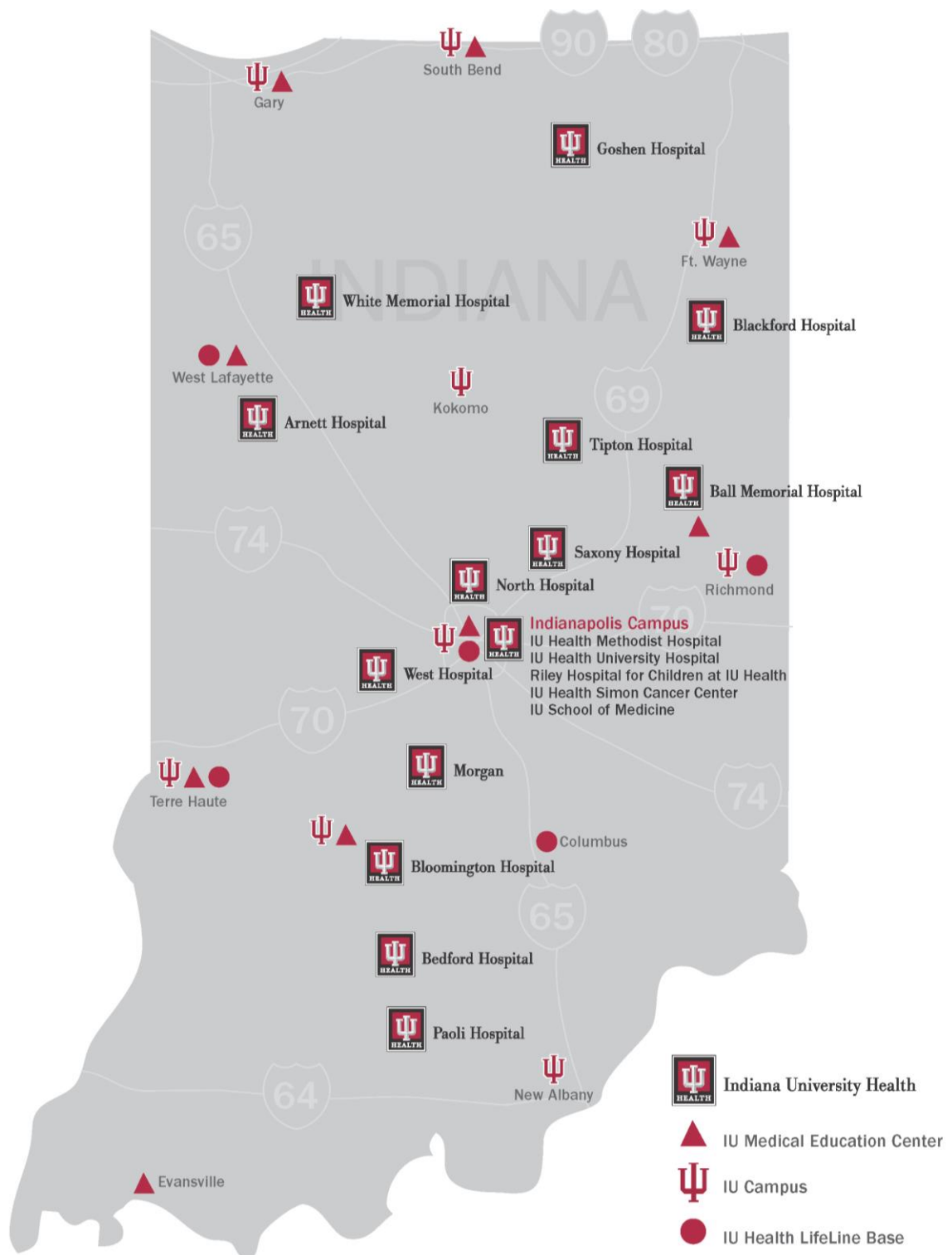


Figure 1. Statewide distribution of IU School of Medicine and IU Health facilities in 2016.

Indiana University Health is a corporation that operates several health centers and hospitals across Indiana, as well as holds a partnership with IU School of Medicine. In Figure 1, “Indiana University Health” locations indicate Indiana University hospitals. The “IU Medical Education Centers” are IUSM campus locations, whereas the “IU Campus” locations include schools in addition to IU School of Medicine. Lastly, “IU Health LifeLine Bases” are emergency helicopter takeoff and landing locations.

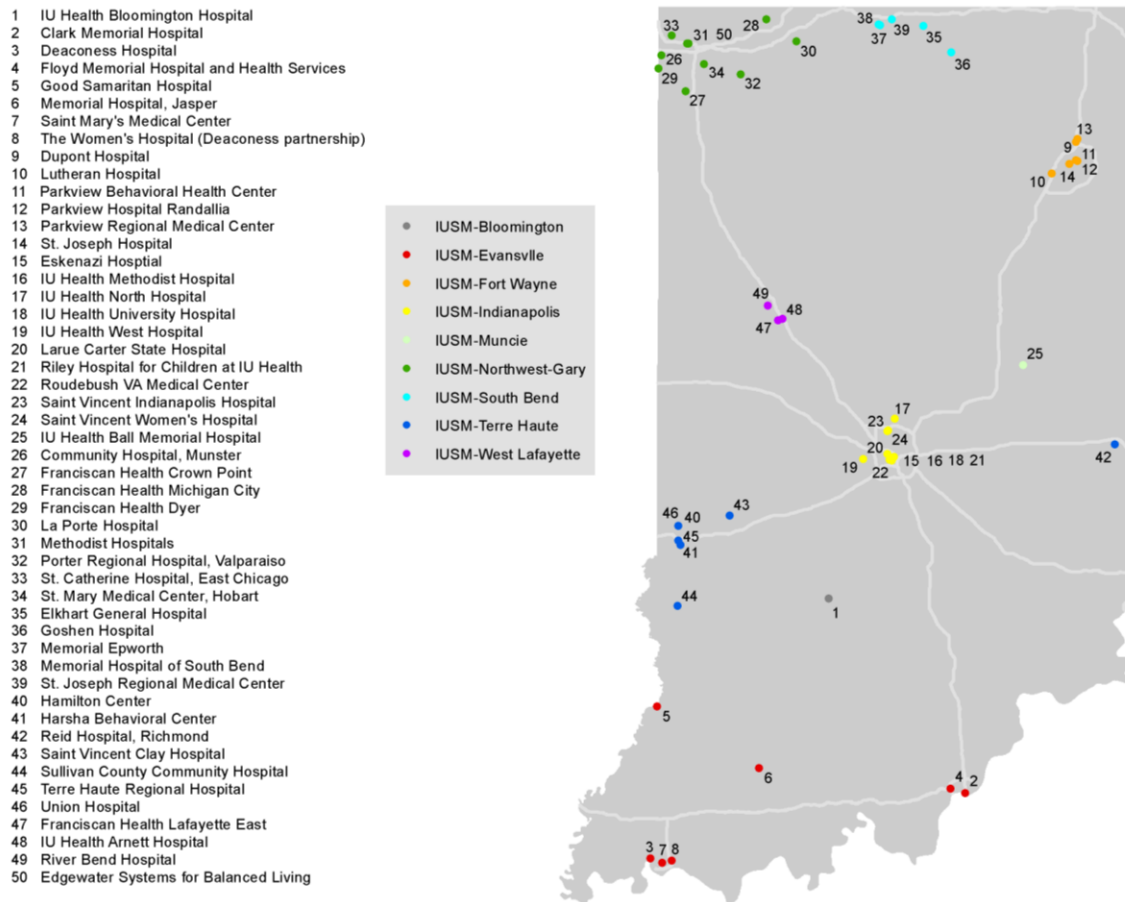


Figure 2. Hospitals and clinical location partners located in Indiana in 2016.

This project represents the first phase of an ongoing collaboration to more deeply integrate geographic technology in the IUSM student data management process. The IU Office of Medical Student Education categorizes student data into distinct departments based on education status: applicants, medical students, and residents/fellows. The first goal was to examine the geographic attributes of different medical student databases maintained by three different offices in the IUSM and to test processes for integrating these data with existing geographic technologies. The three IUSM offices that participated in this collaboration were the Office of Admissions, the Office of Medical Student Education, and the Office of Graduate Medical Education. The second goal was to create a series of prototype maps to visualize spatial and temporal patterns associated with medical school applicants, current medical students, and students engaged in graduate medical education. Specific outcomes included a series of maps that illustrate where applicants, residents and fellows come from, and the relationships between places of student origin, training and specialty area choice dichotomized as primary care vs. non-primary care. Additional outcomes included an evaluation of the geographic components of existing IUSM databases, and a series of recommendations for refining data collection and management processes to more seamlessly facilitate integration of medical student data in geographically-enabled databases for future research.

## BACKGROUND

Geography impacts the decisions that medical students make that in turn help shape our nation's medical landscape. One of the ways that geographic information has been applied in physician workforce research is in the analysis of practice location choice. For example, the locations of medical students' hometowns, undergraduate and graduate institutions, and residency locations have been investigated as predictors of practice location (Wade et al., 2007; Pretorius et al., 2010). Given the shortage of physicians in medically underserved areas (MUAs), several recent studies have focused on examining predictive factors affecting physicians' choice of practice location (Herd et al., 2016; Kost et al., 2014). MUAs are defined as geographic areas that have a shortage of primary care services for local populations. The criteria for defining an MUA is the Index of Medical Underservice (IMU), which scores locations from 0 to 100 (0 representing the most underserved areas and 100 representing the most well served areas). Areas scoring 62.0 or under are considered MUAs. The score is based on four different variables including: primary care physician per 1,000 population, percent of population under the poverty line, infant mortality rate, and population over 65 years old (HRSA, 2016). As of 2017, there were 3,583 designated MUAs in the United States (HRSA, 2016).

### Hometown/Practice Location

The hometown location of medical students has been shown to be a significant predictor of where a student may practice following graduation. For example, in an analysis of 2,487 medical students graduating from Indiana University between 1988–1997, Wade et al. (2007) found that 1,018 out of 1,284 students (73%) practiced in areas

with similar characteristics to their hometown in terms of metro vs. non-metro classification. The conclusions of this study suggested physicians from non-metro hometowns were nearly five times more likely to practice in non-metro locations.

Pretorius et al. (2010) examined the relationship between residency location and practice location at the midcareer point of practicing physicians. Career midpoint was defined as 17-19 years after graduation from medical school. Four categories of place of residence before individuals went on to practice were considered: birth place, high school location, undergraduate location, and medical school location. Two geographic attributes of the physicians' residential location history were significant predictors of career midpoint practice location: birthplace and location of their undergraduate institutions. The authors reported that 84% of individuals had previously lived in their practice area at least once and that 73.9% had lived in the area where they practiced three or more times during their life. The primary conclusion of the study was that prior places of residence were a significant predictor of where physicians practiced at career midpoint.

In a recent study that focused on factors predictive of physicians practicing in rural areas of Australia, Herd et al. (2016) reported that significant predictors included rural vs. urban preference of practice location at time of commencement and the locations where graduates completed their clinical rotations. Data for this study was drawn from questionnaires that were part of a tracking project known as the Medical Schools Outcome Database (MSOD), which provided demographic information and answers to locational questions. The individuals who participated in the survey ( $n = 4,208$ ) were either one or three years post-graduation from medical school. The results showed that students who indicated a preference for rural, rather than urban practice locations at the

time of commencement were significantly more likely to practice in rural locations at one and three years after graduation. Graduates with a rural preference upon commencement were also more likely to choose a rural practice location independent of whether or not they came from a rural background.

### Specialty Area

Not only can geographic characteristics help predict where a medical student might eventually practice, but they can also be predictive of specialty area choice. As previously stated, there has been a significant decrease in the number of medical school graduates practicing in primary care specialties. In 2016, the Association of American Medical Colleges (AAMC) reported that the demand for primary care physicians (PCPs) was increasing while the supply of PCPs was not. The predicted U.S. shortfall could range between 7,300 and 43,100 by 2030 (Dall et al., 2017).

Among the factors contributing to the primary care physician shortage are education debt (Philips et al., 2014) and lifestyle preferences, including a desire for more leisure and family time (Schwartz et al., 2010). Brokaw et al. (2009) concluded that training location was also predictive of specialty choice. Using multivariate logistic regression to analyze data from a ten-year cohort of IUSM graduates, they looked at medical specialties and the campuses of the students' basic sciences courses during their first two years of medical school. Among the regional campuses in the IUSM system, all except the Bloomington campus produced more primary care practitioners than the main campus in Indianapolis. Additionally, all but two regional campuses (Bloomington and Gary) produced a higher number of family care practitioners than the main campus in Indianapolis.



Beginning in 2006, the University of Washington School of Medicine created a program for students who expressed interest in providing care in medically underserved areas. They established an extracurricular opportunity called Underserved Pathways (UP). The program included a variety of components including mentoring, preclinical preceptorships and clerkships at underserved locations, and project-based learning. Kost et al. (2014) conducted a study using data from students that participated in the UP program and those that did not. They reported that UP students matched at a higher rate into primary care specialties compared to those who did not participate in the program. They found that 72.5% of the UP participants matched into the primary care programs, compared to only 48.1% of nonparticipants. More specifically, out of all the primary care graduates, 46% were UP graduates who focused on family medicine. The overall conclusion of the study was that providing targeted training and opportunities for students to participate in medically underserved communities during their education can be an effective strategy for increasing the number of student who chose to practice in primary care specialties.

Canada is also facing the issue of primary care physician shortage. For several years, researchers have been investigating potential causes of and solutions to this problem. Gil et al. (2012) surveyed first, second, and third year medical students at the University of Alberta in Edmonton in May of 2010. The survey included questions about demographics, specialty area preference, and desired future practice location. Results of the study indicated that 43.5% of students with rural backgrounds chose family medicine as their specialty area, as opposed to only 21.5% of those with urban backgrounds. While having a rural background was not the only factor predictive of medical students'

preference for practicing in family medicine, understanding the influence that geography can have on specialty choice could help increase the number of primary care physicians.

## DATA

The IU School of Medicine categorizes student data into three broad areas based on stage of educational career: applicants, medical students, and residents/fellows.

Different offices within the school conduct collection of data during these three different phases. For purposes of the current study, we obtained data from the Office of

Admissions, the Office of Medical Student Education (MSE), and the Office of Graduate Medical Education (GME). Students must fulfill pre-requirements to apply to IUSM.

Once accepted, the Office of Medical Student Education supports students during their four years of medical school. The Office of Graduate Medical Education assists graduates throughout their residency or fellowship (Figure 3).



Figure 3: Medical student education pipeline from applicant to resident/fellow.

While it is necessary for the IUSM administrators to have access to identifiable information including student name, home addresses and university ID numbers, individually identifiable characteristics were only used in the current study when absolutely necessary to link the different databases. Student ID numbers were converted to an arbitrary unique ID. This was done by replacing the Student ID numbers with a random number. Potentially sensitive information included in the IUSM student databases used in this study included student names and street addresses, which were

necessary for geocoding purposes. Where pertinent street address, ZIP code, city/place name, state, county/province, and country fields were retained for generalized mapping purposes. All the data used in the study were stored on systems identified by Indiana University policies as appropriate for internal use. Access to these data were limited to faculty, administrators, and graduate research assistants who completed CITI training, appropriate certification, and who were included on the application approved for this study by the Indiana University Institutional Review Board (IRB# 1711040018).

The IUSM Office of Admissions is the initial point of contact through which prospective students apply to request acceptance into the IU School of Medicine. Data provided by the Admissions Office for this project consisted of a fourteen-page PDF document that included entering student information from 2010 through 2015. Separated by year, these data listed the number of admitted students per county or state of residence at time of admission. Our objective was to examine processes that would be necessary to aggregate these data in order to create maps illustrating the number of Indiana resident students from each county by year, the number of nonresidents from each state by year, and the number of entering students at each of the nine IUSM campuses.

Data provided by the MSE Office included 20 separate Excel files that contained information about individual student campus assignments in their first year of medical school and where they were assigned for subsequent years by class standing. The time period covered by these data included the 2010-11 through 2014-15 academic years. Specific attributes for each student included an arbitrary unique ID number, class year, and campus assignment locations. Each of the 20 files was named according to the

academic and class year. For example, the name “2010-2011 MS3” contained 2010-2011 academic year data for students who were in their third year of medical school.

The Office of Graduate Medical Education (GME) supports residents and fellows while they pursue postgraduate training. Data from the GME office were provided in an Excel file that included 88 worksheets with demographic information for students from the years 2010 through 2016. The fields included: unique ID, program, residency or fellowship, city and country of birth, projected start date, and projected end date. In the current study, these data were used to develop thematic maps that enabled visualization of spatial and temporal patterns of where GME students come from.

Additional data provided by the GME Office indicated for each student their intentions after completing their residency or fellowship in four categories (Figure 4): patient care or clinical practice, fellowship or additional specialty training, academic position (teaching and/or research), or other (e.g., military or undecided). The objective in the current study was to use these data to create thematic maps that illustrate the total number and proportion of students who intended to pursue primary care vs. non-primary care specialties at each campus across the different class years.

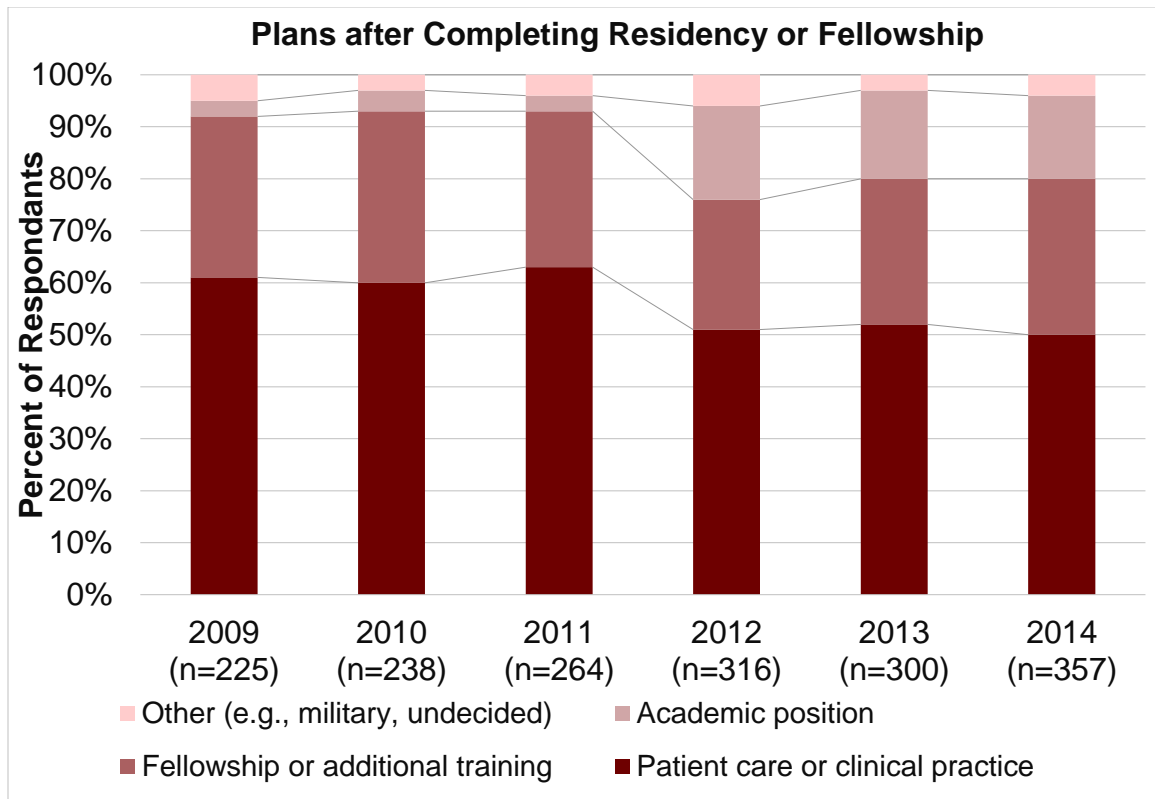


Figure 4: Residents and fellows plans after completing their residency or fellowship training at IUSM based on GME databases.

## METHODS

While the data processing and cartographic design methods employed in this study are fundamental tasks from a GIS perspective, the purpose was to examine and document the processes necessary to integrate data from three different IUSM student service offices into a common GIS environment. One goal of the study was to examine barriers to efficient conversion and geocoding that need to be considered in support of a longer-term vision of creating a system that facilitates advanced health workforce research incorporating spatial analysis. A second goal of the study was to develop a series of prototype maps that would serve as starting points for discussions about the of design, content, and desired medium for presentation of medical student data (e.g., hardcopy, web, digital presentation, or animation).

Basic data processing tasks had to be implement throughout this study to integrate data from the different IUSM departments. For example, several PDF documents needed to be converted into spreadsheets to import them into a GIS environment. Most of the data was cleaned in Excel, which included searching for and correcting spelling errors as well as removing unnecessary punctuation. Concatenating data fields was also implemented in order to enable linking some files together. For example, to create the IUSM: Breakdown by Campus, 2010-2014 map, we were given an Excel file, (2010-11 MS1) that listed the campus assignment for 339 students who were in their first year of medical school. The excel file included “2010-2011 Class Year”, “University ID”, and “First Year Campus” location. These data needed to be joined with the 2014 specialty area file in order to determine the practice specialty preference for each student.

Since consistent ID numbers were not available for each student in both databases, the join had to be implemented using student names. However, in both files, the names were stored in different formats. In the 2010-2011 MS1 file, the names were displayed as “Last Name, First Name Middle Initial.” In the 2014 specialty area database, names were presented as, “Last Name, First Name Middle Name,” and in some cases middle names were not given. This inconsistency made what could be an easy, automated process more time consuming and required manual review. Special characters included in name fields also contributed to some errors in matching (e.g. Le-Ann). Special characters were either removed or standardized through manual review before completing the concatenation.

The concatenation process was implemented in ArcMap using the Python code shown in Figure 5. Four new text fields were created for ‘first name’, ‘last name’, ‘middle name’, and the new concatenated field ‘last name, first name’ in both tables. The Python code was used in ArcGIS’s Field Calculator to separate the original ‘first, middle, and last name’ fields, and eventually the first and last names were joined together in one field within each table.



```
Last Name: Field Calculator > Python.string >  
!Graduate!.split(",")[0]  
  
First Name: Field Calculator > Python.string >  
!Graduate!.split(",")[1].split(" ")[1]  
  
Middle Name: Field Calculator > Python.string >  
!Graduate!.split(" ")[2]  
  
Last Name, First Name: Field Calculator > Python.string >  
!Last_Name! + "@" + !First_Name!
```

Figure 5: The Python code used to concatenate first and last name fields into one field.

After this cleaning phase, the new concatenated name fields were used to join both tables. Next, separate files were created for students that matched into a Primary Care (PC) and Non-Primary Care (NPC) specialty. These data were summarized based on the number of PC and NPC individuals at each campus location. The summary data were then joined with campus location shapefiles and the information was symbolized using separate pie charts showing the percent PC and NPC for each campus.

All maps were designed with a minimalist aesthetic for practical reasons. During the internship phase of the project, iterative drafts of different maps were presented to IUSM administrators to discuss design and content choices. Requests by administrators to share these maps in organizational meetings led to the recognition that both the contexts in which the maps would be presented and the mode of presentation could vary (e.g.,

hardcopy reports, projected graphics, etc.). Adaptability to diverse audiences with varied backgrounds was another important consideration (for example, internal meetings, external evaluators, or legislators). Thus, a minimalist design strategy that allowed for customization depending upon context and also for scalability for inclusion into a variety of document or media formats was adopted.

Twelve maps were developed from the medical student data provided by the Office of Admissions including both county and national level variations (Figures 6a – 6f, 7a – 7f). Basic data processing tasks included converting PDF files to text format for import into the GIS and creating aggregate counts of students by county and state using standard database summary operations. While not part of the primary mapping goals, similar operations were performed at the place level (city, town or census designated place) to evaluate the potential for mapping at finer levels of geographic specificity. City and county names from medical student data were joined to geographic files from the US Census Bureau, while documenting the frequency and nature of geocoding errors resulting during the process.

Six thematic maps were developed from data supplied by the IUSM Medical Student Education (MSE) Office. As mentioned, the relevant fields in the MSE data included class year, current campus, 1<sup>st</sup> year campus assignment, and planned specialty area. Following similar pre-processing procedures implemented with the admissions data, five maps were created to display the number of students by each class year at each of the nine IUSM campuses for the 2010-11 through 2014-15 academic years (Figures 8a – 8e). A sixth map depicting medical students graduating from the years 2010 through 2014

matching into primary care (PC) and non-primary care (NPC) residencies was created based on 1<sup>st</sup>-year campus assignment location (Figure 9).

The Graduate Medical Education Office supplied documentation on where the graduate students originated from including city/place, state/province, and country. For the GME map (Figure 10), we focused on the “Country” field, but we also explored finer levels of location detail by investigating international sources of place names for future research.

Finally, while this study did not focus on presentation of individual-level data, one of the long-term goals of IUSM administrators is to be able to examine career paths of their students over the life course, from place of birth, through their undergraduate and medical education, residencies and graduate training, to their eventual place(s) of practice or other forms of contribution to the field of medicine. This closely ties into the conceptual framework of time geography (Hägerstrand, 1970), which is becoming increasingly practical to implement given the growing accumulation of individual-level data attributed with place and time (Miller, 2005; Kwan and Neutens, 2014). Thus, a final outcome of the project was a hypothetical map of IUSM student career paths that was intended to spark further discussion and consideration of advancing GIS integration in IUSM systems.

## RESULTS

### Map Narratives

The goals of this project included examining geographic attributes of three different IUSM student databases and integrating these data with current geographic technologies, as well as creating prototype maps to visualize patterns using the medical student data. These maps are briefly described below.

#### Admissions Maps

The goal for creating the admissions maps was to visualize in-state applicants by county of origin, and at the national level by state. The original data was in pdf format and displayed a count of the number of each applicant by Indiana county or state, by year. Two series of maps were created: one focusing on the number of in-state students by Indiana County and the second series focusing on the number of students from each US state.

#### Indiana County Maps

Six maps showing applicants by Indiana county for the years 2010 – 2016 were created (Figure 6a-6f). Four out of six of the ‘Applicants by County’ maps indicate that the largest number of applicants come from Marion County, which is where the main IUSM campus is located in Indianapolis. Over the years, IUSM has not only increase their student enrollment size, but also recruited more students from within the state. Being a state supported medical school, IUSM favors Indiana applicants over non-residents. In addition, some regional campuses are able to fill a sizeable portion of their entering class with students from the surrounding campus region.

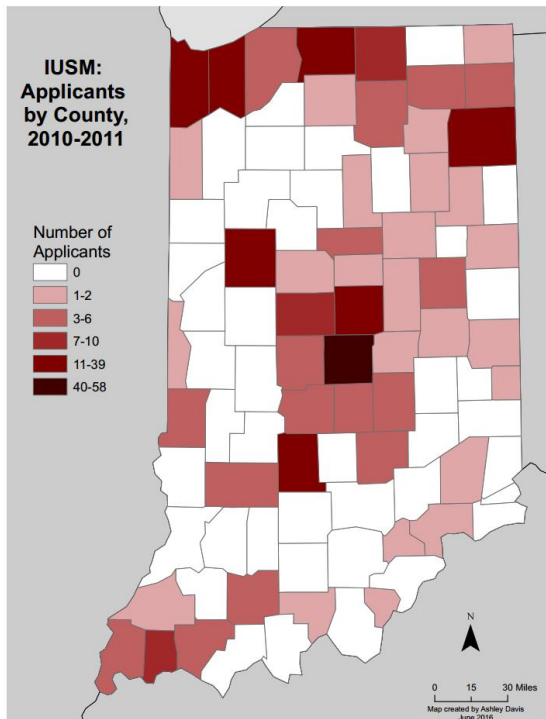


Figure 6a.

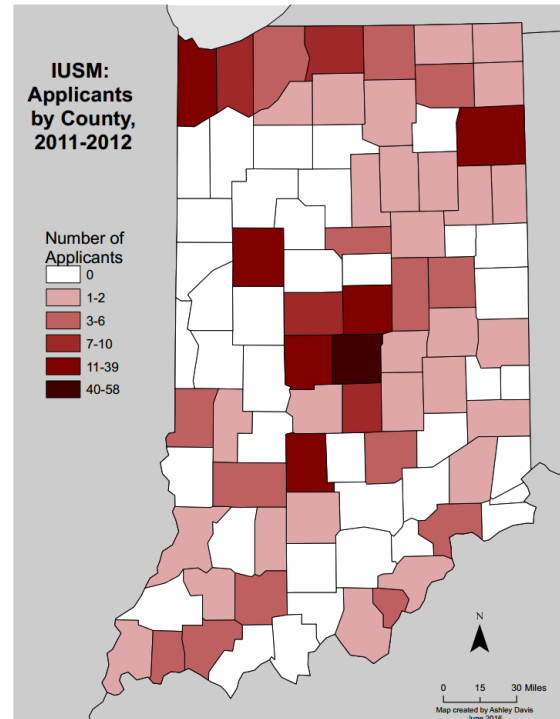


Figure 6b.

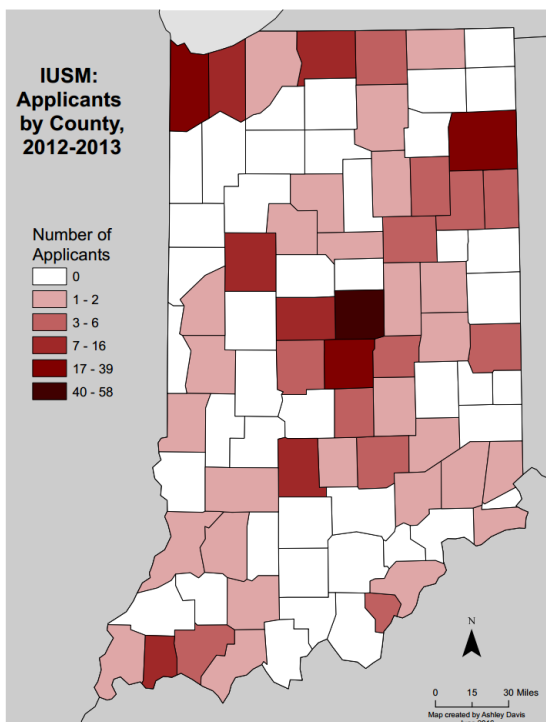


Figure 6c.

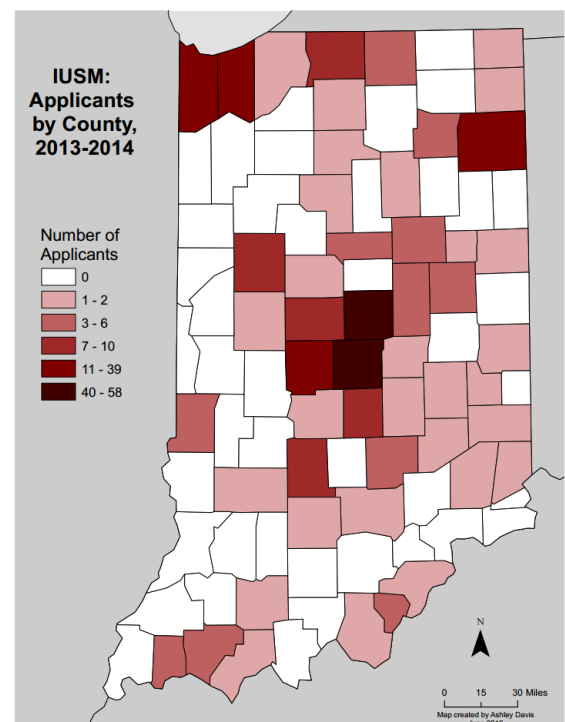


Figure 6d.

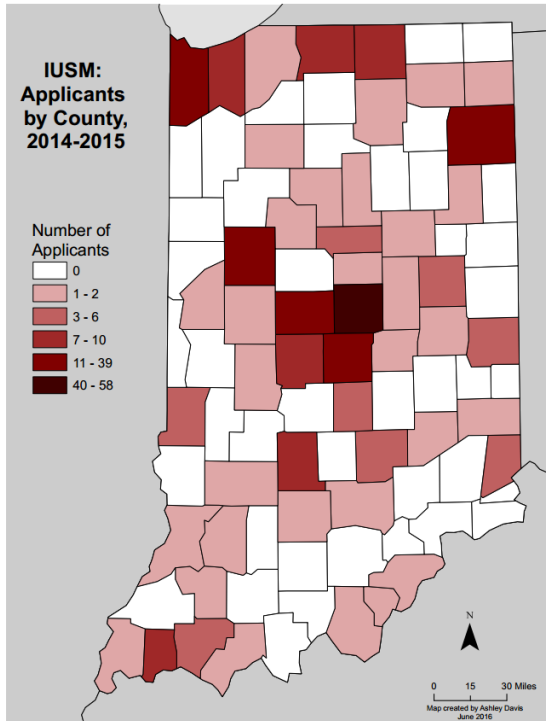


Figure 6e.

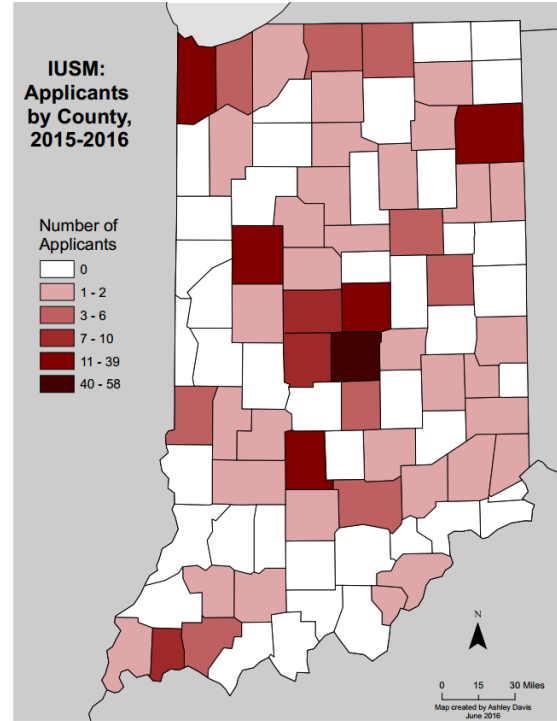


Figure 6f.

Figures 6a – 6f: Admissions maps identifying the number of IUSM applicants from Indiana counties during the 2010 - 2015 academic years.

### United States Maps

The United States maps (Figures 7a - 7f) were created with the same goal as the Indiana maps, to show counts by state from which IUSM applicants originated. As expected, Indiana had the highest number of applicants in all years. Indiana applicants ranged between 263 to 289 applicants per year from 2010-2016. There was a significant gap between counts of applicants from Indiana and the next highest state - California, which had 19 applicants. California, the most populous state in the nation, was consistently in the top three states with the highest number of applicants outside of Indiana for the data period examined. Other states that had relatively high numbers of applicants included Michigan, ranging from 6-7 applicants, Ohio, ranging from 3-6

applicants, and Illinois, with 3 to 13 applicants. Since these states all border Indiana, it is not surprising that they had a relatively high number of applicants. Some of the outliers that appeared during only one year included Missouri, with four applicants in 2013-2014 (Figure 7c), and New York and Utah, both with five applicants in the 2015-2016 academic year (Figure 7f).

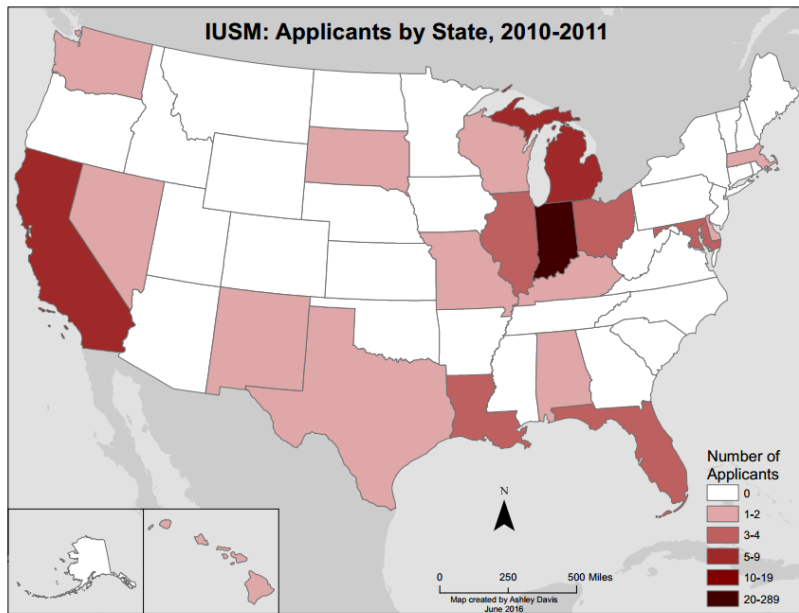


Figure 7a.

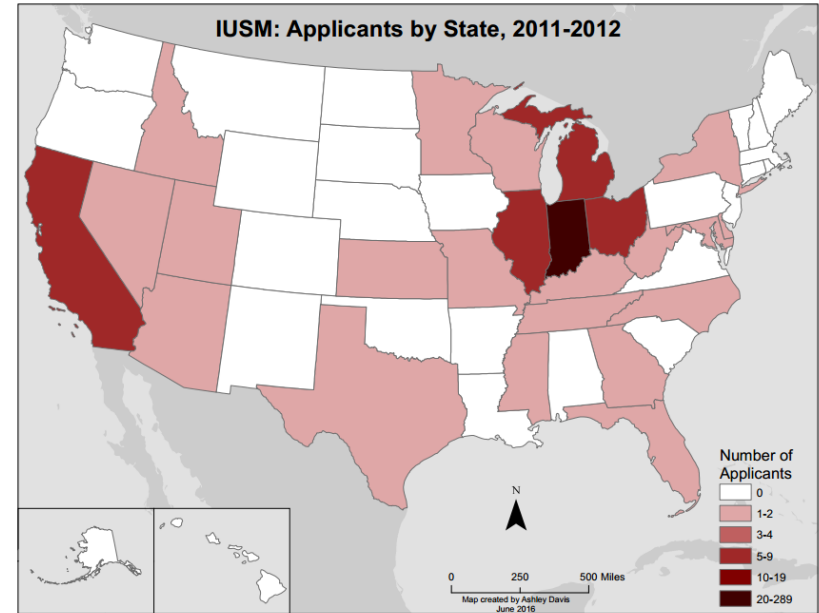


Figure 7b.



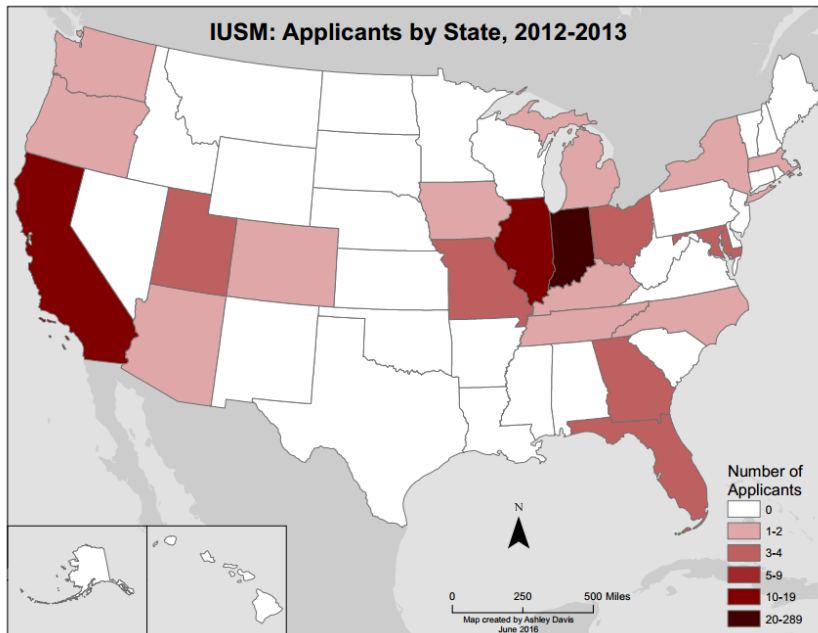


Figure 7c.

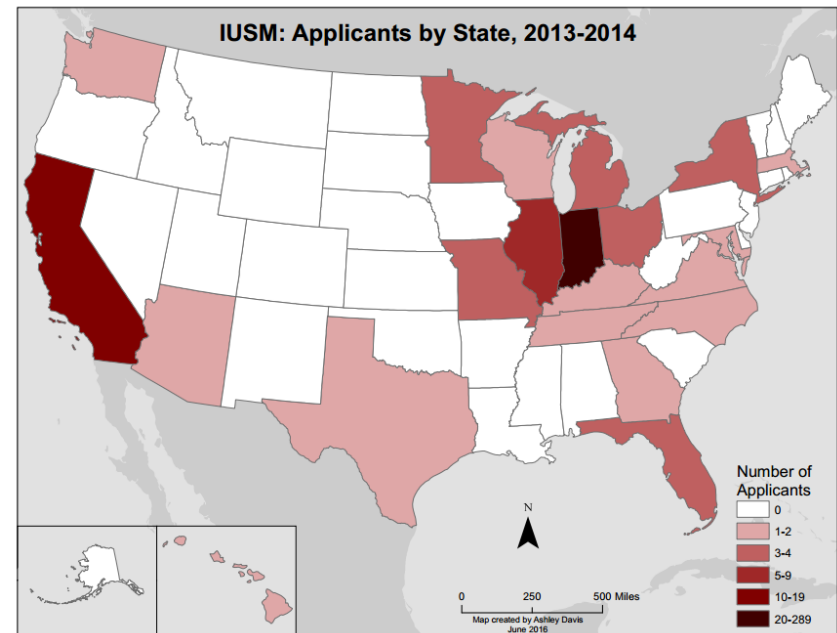


Figure 7d.

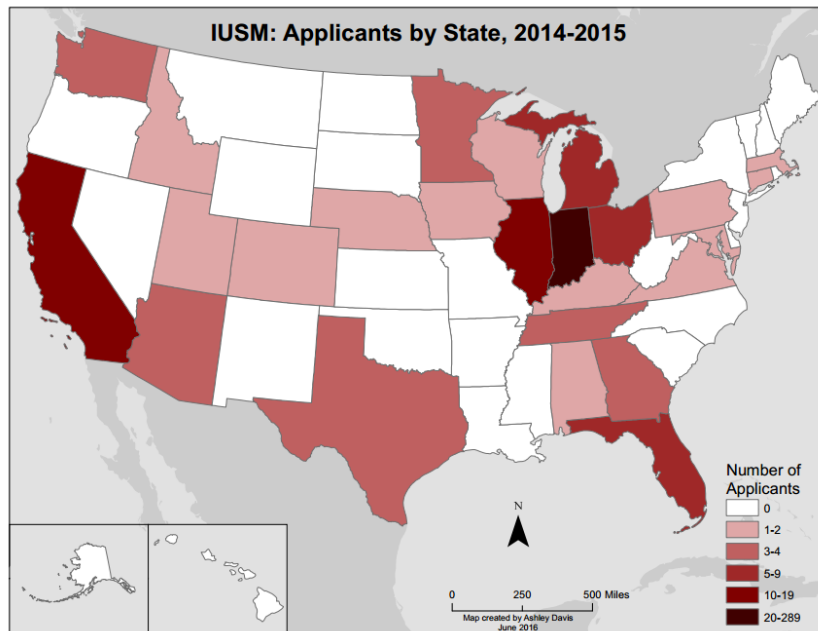


Figure 7e.

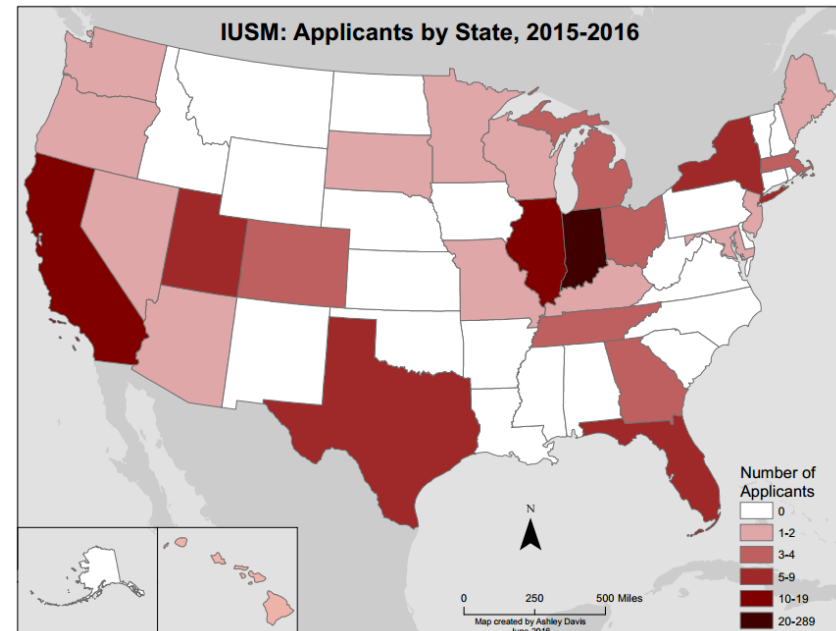


Figure 7f.

Figure 7a -7f: Admissions maps displaying the number of IUSM applicants by state during the 2010 – 2015 academic years.

### Medical Student Education Maps

These maps were created to illustrate student assignment at each IUSM campus using bar graphs representing the number of students enrolled at a campus by individual class year. In the legend, 'n' represents the total number of students in the class year (Figures 8a – 8e).

One of the most noticeable trends in these maps are the differences in majority class year among the different campuses. While the main IUSM campus in Indianapolis has the highest number of students among all class years, it is important to note that before the 2012-2013 academic year (Figure 8c), curricular policies required almost all medical students to come to the Indianapolis campus for their final two years. However, the policies changed so that a student could be assigned to a regional campus for their first two years and then return to the Indianapolis campus for their final two years, or they could choose to complete their last two years at the original regional campus.

An additional map developed from the MSE data presents student specialty area matching by campus location (Figure 9). This map was developed using 2010 first year campus location and 2014 specialty area data. Before creating the map, there was an assumption that the students assigned to the Indianapolis campus during their first year were more likely to match into non-primary care (NPC) specialty areas compared to the other regional campuses. While Indianapolis had a 56% NPC match rate, the Fort Wayne campus had the same rate, and the Bloomington campus had a higher NPC rate at 67%. The Terre Haute, Muncie, Gary, and South Bend campuses were comparable with only a 9-12% difference. However, the Lafayette and Evansville campuses had a lower rate of NPC matches at 30% and 27% respectively.

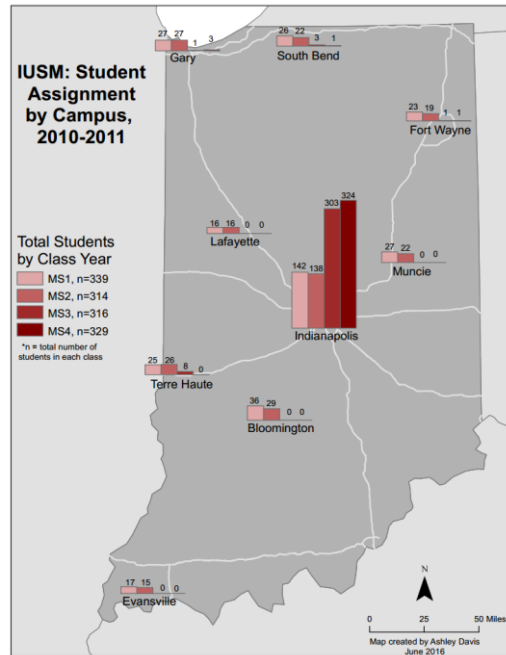


Figure 8a.

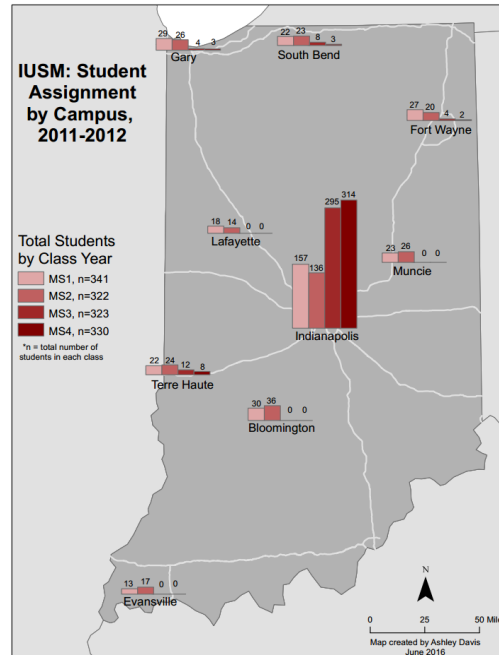


Figure 8b.

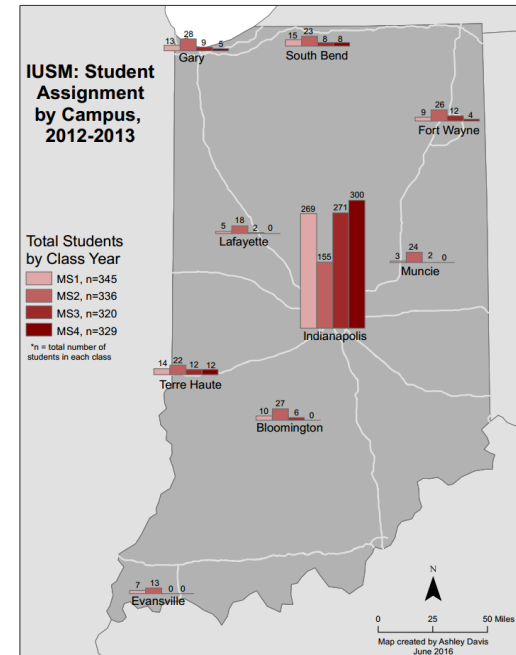


Figure 8c.

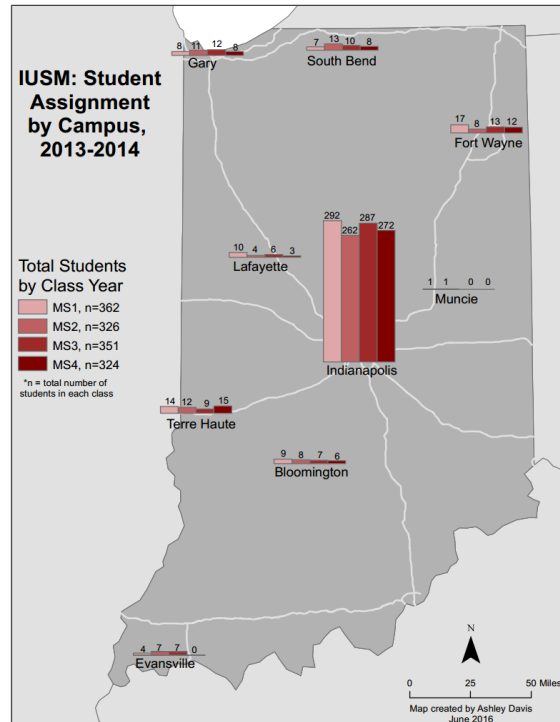


Figure 8d.

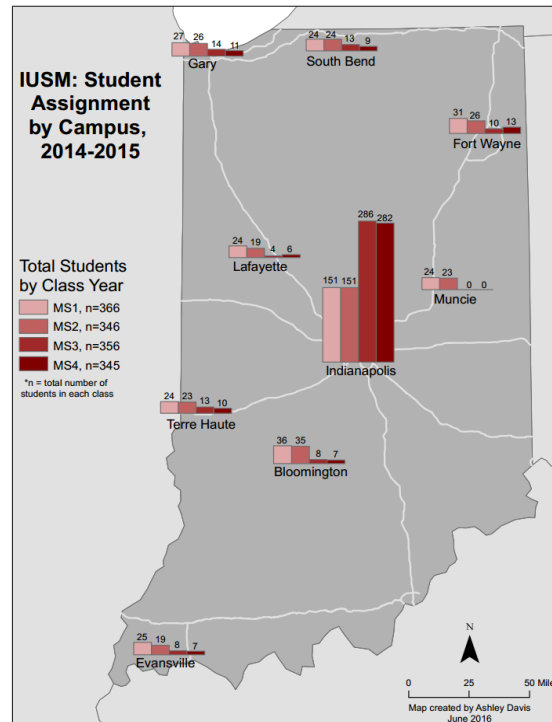


Figure 8e.

Figure 8a – 8e: Medical Student Education (MSE) maps illustrating the total number of students by class year at each campus during 2010 – 2014 academic years

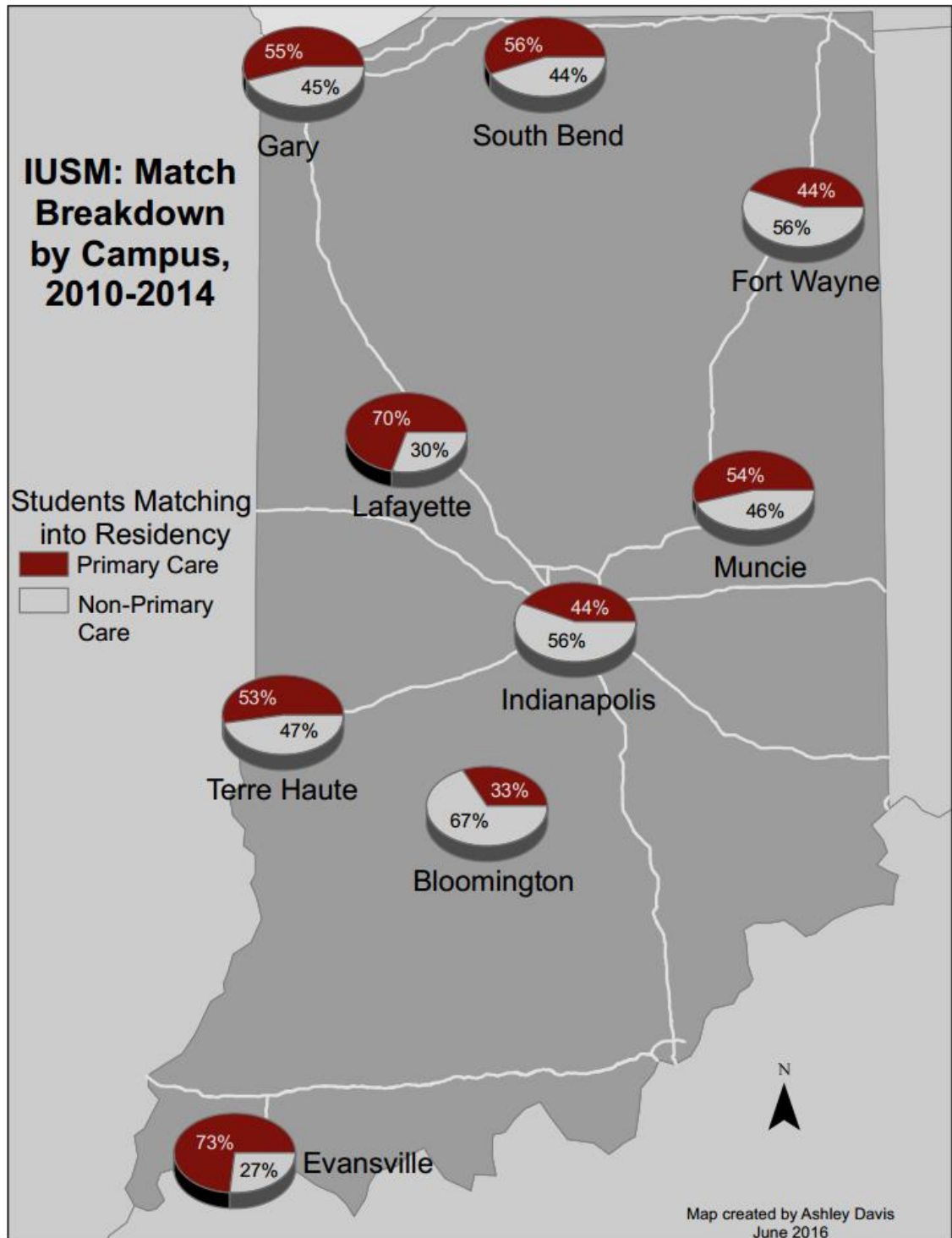


Figure 9: Percent of medical students matching into either primary care or non-primary care residency programs by first year campus location.

### Graduate Medical Education World Map

The graduate medical education world map (Figure 10) developed using data from the GME Office identifies the countries from which IUSM residents and fellows originated summarized over the 2010-2016 academic years. As expected, the vast majority of applicants accepted into the residency program were from the US, with 1,526 residents, followed by India (95), Pakistan (42), China (35), and Canada (34).

## IUSM: Graduates entering Residency by Country of Origin

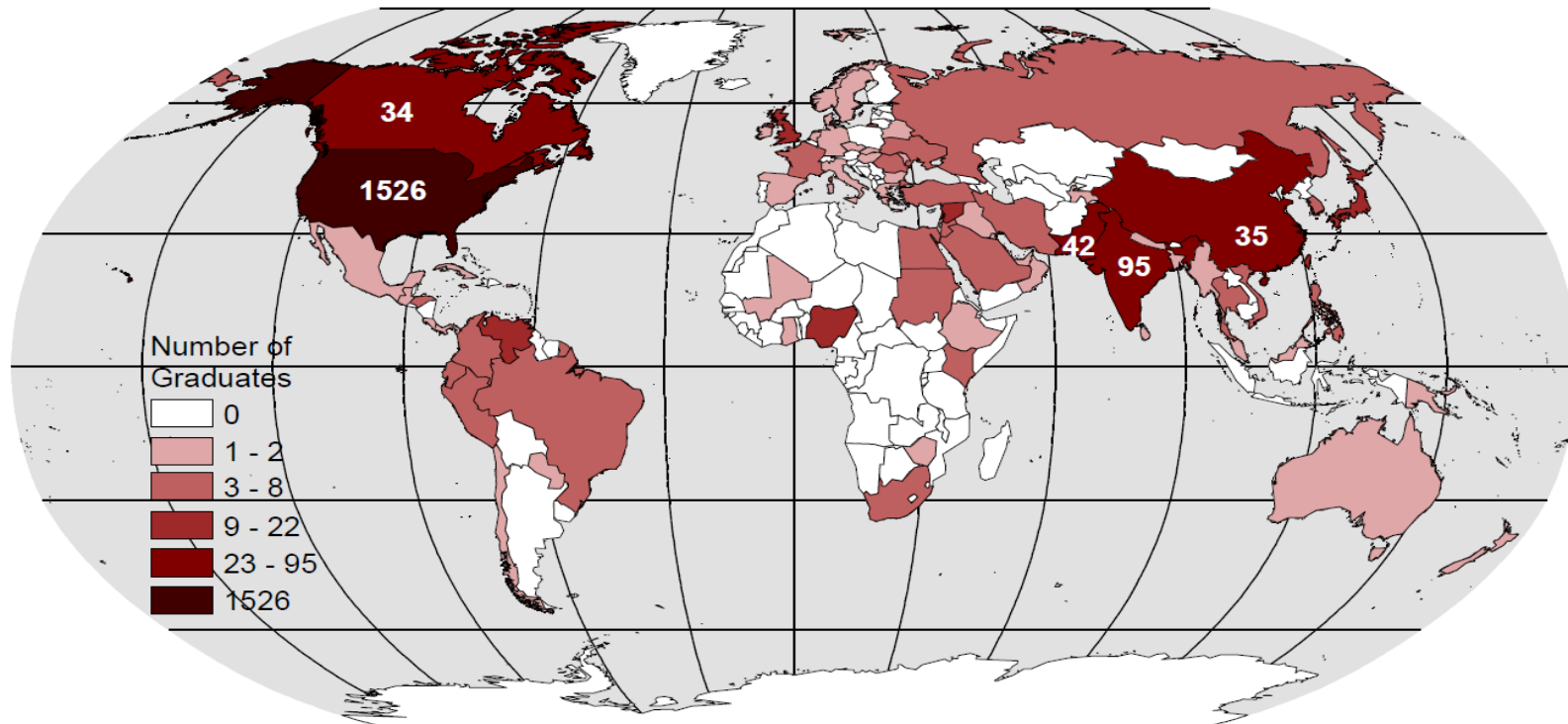


Figure 10: The Graduate Medical Education (GME) world map, depicting the country of origin for graduates entering the residency or fellowship program at IUSM.



## DISCUSSION

The primary goal of producing an initial series of maps that show the spatial patterns of IUSM student data was achieved but working with the data we received from the various offices presented challenges that highlight opportunities for more efficient data collection and standardization that would streamline integration of these data with spatial technologies.

A primary obstacle was the lack of consistency in the unique identifiers for individual students in the different datasets. While each set of data had unique identifiers, none of these matched the other datasets. This inhibited efficient linking of data and following students through their educational paths. Another hurdle encountered during included spelling conventions. In several databases, there were many misspellings of Indianapolis, Indiana, rendering it useless to run counts on the data without sorting through and cleaning the spelling errors first. Another difficulty was inconsistency in the data formats from the different IUSM offices. For example, data from the Admissions Office were provided in PDF format, which required reformatting in order to integrate with the other databases.

## CONCLUSIONS

This study was the first attempt at exploring the potential to link multiple databases from three different offices at IUSM and integrate the data in GIS-compatible formats. As a result, several recommendations emerged that would make future efforts to be more efficient and facilitate deeper analysis into physician workforce research.

Overall, establishing data standards across all IUSM offices that collect and process medical student data would be key a first step to make the process more efficient and facilitate easier integration with geographic technologies.

Standardization of data file formats is one key step that can be taken. While PDFs work well for the presentation of information, this format does not facilitate easy integration into tabular formats. This impedes efficiency in data processing and takes more time overall. Standardizing formats using formats such as csv, Excel, or other common database formats is recommended.

Next, there should be attribute domains in place to prevent or at reduce data entry errors. Attribute domains are the set values for a data category. For example, for a category named “City,” there would be a drop-down list of city names to choose from, rather than requiring an individual to manually enter in the name of each city. This would prevent not only common spelling errors, but also alleviate non-standardized names (e.g. La Grange County vs. LaGrange County). Another suggestion is to use the same unique identifier for individuals throughout their time at IUSM. Doing so would allow further research that could track students over time. Unfortunately, were unable to do this because the unique IDs for students provided by each department were not consistent. Lastly, staff training and education would promote higher quality data and be invaluable

to creating efficient and consistent data entry processes to facilitate research applications of the data.

There are a number of ways that the continuation of using GIS as a tool to facilitate medical student data at IUSM. One of the future uses of this technology would be to use GIS to explore research on the healthcare workforce. Workforce research would take this project a step further and use GIS to study medical students post-graduation, and into their practice. Another way in which this project could impact future studies is by communicating data in such a way that encourages administrators to ask new questions and enhance communication amongst the departments of IUSM, bringing them together. Lastly, these tools could be used to educate different audiences who influence education. One group that could use this information would be legislators, who support medical education and IUSM. Another group who could use this data would be recruiters. They could use the data to find where there may be shortages of doctors or even applicants, and then recruit students from those areas, with the idea that those students will eventually return and enhance their community.

## REFERENCES

- Brokaw, J. J., Mandzuk, C. A., Wade, M. E., Deal, D. W., Johnson, M. T., White, G. W., ... & Zollinger, T. W. (2009). The influence of regional basic science campuses on medical students' choice of specialty and practice location: a historical cohort study. *BMC Medical Education*, 9(1), 29.
- Dall, T., West, T., Chakrabarti, R., & Iacobucci, W. (2017). The complexities of physician supply and demand 2017 update: projections from 2015 to 2030. *Washington: Association of American Medical Colleges*.
- Gill, H., McLeod, S., Duerksen, K., & Szafran, O. (2012). Factors influencing medical students' choice of family medicine Effects of rural versus urban background. *Canadian Family Physician*, 58(11), e649-e657.
- Hägerstrand, T. (1970). What About People in Regional Science? *Papers of the Regional Science Association* **24**, 7–21.
- Herd, M. S., Bulsara, M. K., Jones, M. P., & Mak, D. B. (2017). Preferred practice location at medical school commencement strongly determines graduates' rural preferences and work locations. *Australian Journal of Rural Health*, 25(1), 15-21.
- HRSA (Health Research and Services Administration). (2016) Medically underserved areas and populations (MUA/Ps). Retrieved from: <https://bhw.hrsa.gov/shortage-designation/muap>
- HRSA (Health Research and Services Administration). (2017). Medically underserved areas/populations (MUA/P) state summary of designation MUA/P [Data file]. Retrieved from [https://ersrs.hrsa.gov/ReportServer?/HGDW\\_Reports/BCD\\_MUA/](https://ersrs.hrsa.gov/ReportServer?/HGDW_Reports/BCD_MUA/)

Indiana University School of Medicine. (2016). Fact sheet 2016-17. Retrieved from:

<https://mednet.iu.edu/MasterDocLibrary/IUSM-Fact-Sheet.pdf>

Kost, A., Benedict, J., Andrilla, C. H. A., Osborn, J., & Dobie, S. A. (2014). Primary care residency choice and participation in an extracurricular longitudinal medical school program to promote practice with medically underserved populations. *Academic Medicine*, 89(1), 162-168.

Kwan, M. P., & Neutens, T. (2014). Space-time research in GIScience. *International Journal of Geographical Information Science*, 28(5), 851-854.

Miller, Harvey J. "A measurement theory for time geography." *Geographical analysis* 37.1 (2005): 17-45.

Office of International Affairs. (2016). International enrollment report September 2016.

Retrieved from: <http://international.iupui.edu/doc/about/2016-intl-student-enrollment.pdf>

Phillips, J. P., Petterson, S. M., Bazemore, A. W., & Phillips, R. L. (2014). A retrospective analysis of the relationship between medical student debt and primary care practice in the United States. *The Annals of Family Medicine*, 12(6), 542-549.

Pretorius, R. W., Lichter, M. I., Okazaki, G., & Sellick Jr, J. A. (2010). Where do they come from and where do they go: implications of geographic origins of medical students. *Academic Medicine*, 85(10), S17-S20.

Schwartz, M. D., Durning, S., Linzer, M., & Hauer, K. E. (2011). Changes in medical students' views of internal medicine careers from 1990 to 2007. *Archives of internal medicine*, 171(8), 744-749.

United States Census Bureau. (2016). Annual estimates of resident population: April 1, 2010 to July 1 2016. Retrieved from <https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk>

Wade, M. E., Brokaw, J. J., Zollinger, T. W., Wilson, J. S., Springer, J. R., Deal, D. W., White, G. W., Barclay, J. C., & Holloway, A. M. (2007). Influence of hometown on family physicians' choice to practice in rural settings. *Family Medicine*, 39(4), 248-254.

## CURRICULUM VITAE

Ashley Michelle Davis

### **EDUCATION**

Master of Science in Geographic Information Science, Indiana University Purdue

University – Indianapolis (August 2014 – April 2019)

Bachelor of Science in Social Studies Education, Indiana University Purdue University –

Indianapolis, (August 2009 - May 2013)

### **CONFERENCES ATTENDED**

American Association of Geographers Annual Meeting, San Francisco, CA, Poster

Presentation (2016)

American Association of Geographers Annual Meeting, Chicago, IL, Poster Presentation

(2015)

### **PROFESSIONAL EXPERIENCE**

#### **Hillsborough County, Public Works Department, Tampa, Florida**

December 2017 – Present

GIS Analyst

- Perform complex analysis of GIS data based on specified criteria and generate reports supporting other County departments
- Maintain GIS databases and design schema
- Design maps and databases; analyze reports; and makes recommendations

#### **Hillsborough County, Public Works Department, Tampa, Florida**

April 2017 – December 2017

GIS Technician

- Create, update, and digitize GIS layers to maintain databases
- Produce weekly reports and maps based on specified GIS criteria
- Assist other County departments with maps and GIS data

**Indiana Department of Transportation (INDOT), Indianapolis, Indiana**

November 2016 – February 2017

GIS Analyst

- Produce county funding maps for Indiana Governor's Office, aeronautical maps, bike suitability maps, and a variety of road maps
- Maintain and perform spatial analyses on roadway and bridge data
- Create new feature classes and databases to use in ArcMap

**Simon Property Group, Indianapolis, Indiana**

March 2016 – November 2016

Market Research Intern

- Support the Research Department's collection, management, and analysis of Simon spatial and non-spatial data
- Produce data layers, maps, demographic reports and sales reports, using spatial analysis procedures and Geographic Information Systems (GIS) technology
- Gather and integrate information from various internal and external sources to fulfill objectives of the department of real estate research to support leasing, executive decision making, and development initiatives of the company

**Indiana University Purdue University - Indianapolis, Indianapolis, Indiana**

August 2014 - June 2016

Research Intern & Graduate Research Assistant



- Analyze GIS data to identify spatial relationships and display data via maps, graphs, and/or tables
- Review/validate internal and external data and data sources for accuracy, usefulness, quality control, and/or completeness of documentation
- Produce data layers, maps, and reports, using spatial analysis procedures and GIS technology (i.e. Queries by location, attributes, etc.)
- Write and present reports of research findings
- Gather and compile geographic data from sources including field observations, censuses, satellite imagery, and existing maps